



AWS Guide de décision

Choisir un service d'apprentissage AWS automatique



Choisir un service d'apprentissage AWS automatique: AWS Guide de décision

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Les marques commerciales et la présentation commerciale d'Amazon ne peuvent pas être utilisées en relation avec un produit ou un service extérieur à Amazon, d'une manière susceptible d'entraîner une confusion chez les clients, ou d'une manière qui dénigre ou discrédite Amazon. Toutes les autres marques commerciales qui ne sont pas la propriété d'Amazon sont la propriété de leurs propriétaires respectifs, qui peuvent ou non être affiliés ou connectés à Amazon, ou sponsorisés par Amazon.

Table of Contents

Guide de décision	1
Introduction	1
Comprendre	2
Tenez compte	4
Choix	7
Utiliser	10
Explorez	23
Ressources	23
Historique de la documentation	25
.....	xxvi

Choisir un service d'apprentissage AWS automatique

Choisissez les services et frameworks de machine learning adaptés à votre travail

Objectif	Aidez à déterminer les services AWS ML les mieux adaptés à vos besoins.
Dernière mise à jour	3 mai 2024
Services couverts	<ul style="list-style-type: none">• Amazon Augmented AI• Amazon CodeGuru• Amazon Comprehend• Amazon DevOps Guru• Amazon Forecast• Amazon Kendra• Amazon Lex• Amazon Personalize• Amazon Polly• Amazon Rekognition• Amazon SageMaker AI• Amazon Textract• Amazon Transcribe• Amazon Translate

Introduction

Dans sa forme la plus élémentaire, l'apprentissage automatique (ML) est conçu pour fournir des outils et des services numériques permettant d'apprendre à partir des données, d'identifier des modèles, de faire des prédictions, puis d'agir en fonction de ces prédictions. Presque tous les systèmes d'intelligence artificielle (IA) actuels sont créés à l'aide du ML. Le ML utilise de grandes quantités de données pour créer et valider une logique de décision. Cette logique de décision constitue la base du modèle d'IA.

Les scénarios dans lesquels les services d'apprentissage AWS automatique peuvent être appliqués incluent :

- Cas d'utilisation spécifiques : les services d'apprentissage AWS automatique peuvent prendre en charge vos cas d'utilisation basés sur l'IA grâce à une large gamme d'algorithmes, de modèles et de solutions prédéfinis pour les cas d'utilisation et les secteurs d'activité courants. Vous avez le choix entre 23 services préformés, dont Amazon Personalize, Amazon Kendra et Amazon Monitron.
- Personnalisation et évolutivité de l'apprentissage automatique : Amazon SageMaker AI est conçu pour vous aider à créer, former et déployer des modèles de machine learning adaptés à tous les cas d'utilisation. Vous pouvez créer les vôtres ou accéder à des modèles de base open source AWS via Amazon SageMaker AI et Amazon Bedrock.
- Accès à une infrastructure spécialisée : utilisez les frameworks et l'infrastructure de machine learning fournis AWS lorsque vous avez besoin d'une flexibilité et d'un contrôle accrus sur vos flux de travail d'apprentissage automatique et que vous êtes prêt à gérer vous-même l'infrastructure et les ressources sous-jacentes.

Ce guide de décision vous aidera à poser les bonnes questions, à évaluer vos critères et votre problème commercial, et à déterminer les services les mieux adaptés à vos besoins.

Comprendre

Alors que les organisations continuent d'adopter les technologies d'IA et de ML, l'importance de comprendre et de choisir parmi les services de AWS ML constitue un défi permanent.

AWS fournit une gamme de services de machine learning conçus pour aider les entreprises à créer, former et déployer des modèles de machine learning plus rapidement et plus facilement. Ces services peuvent être utilisés pour résoudre un large éventail de problèmes commerciaux tels que la prévision du taux de désabonnement des clients, la détection des fraudes et la reconnaissance vocale et d'image.

What is it?



Artificial intelligence (AI)

Any technique that enables computers to mimic human intelligence using logic, if-then statements, and machine learning



Machine learning (ML)

A subset of AI that uses machines to search for patterns in data to build logic models automatically



Classification AI and Predictive AI

A subset of ML that recognizes patterns to identify something (Classification AI) or predicts future trends based on statistical patterns and historical data (Predictive AI)



Generative AI

A subset of DL that can create new content and ideas powered by large, pretrained models called foundation models (FMs)

Avant d'approfondir les services de AWS ML, examinons la relation entre l'IA et le ML.

- À un niveau élevé, l'intelligence artificielle est un moyen de décrire tout système capable de reproduire des tâches qui nécessitaient auparavant l'intelligence humaine. La plupart des cas d'utilisation de l'IA visent un résultat probabiliste, c'est-à-dire la prise d'une prédiction ou d'une décision avec un degré élevé de certitude, similaire au jugement humain.
- Presque tous les systèmes d'IA actuels sont créés à l'aide de l'apprentissage automatique. Le ML utilise de grandes quantités de données pour créer et valider une logique de décision, connue sous le nom de modèle.
- L'IA de classification est un sous-ensemble du ML qui reconnaît les modèles pour identifier quelque chose. L'IA prédictive est un sous-ensemble du ML qui prédit les tendances futures sur la base de modèles statistiques et de données historiques.
- Enfin, l'IA générative est un sous-ensemble de l'apprentissage profond qui permet de créer de nouveaux contenus et idées, tels que des conversations, des histoires, des images, des vidéos et de la musique. L'IA générative est alimentée par de très grands modèles préentraînés sur de vastes corpus de données, appelés modèles de base ou. FMs [Amazon Bedrock](#) est un service entièrement géré qui propose un choix de services très performants FMs pour créer et faire évoluer

des applications d'IA générative. [Amazon Q Developer](#) et [Amazon Q Business](#) sont des assistants basés sur l'IA générative destinés à des cas d'utilisation spécifiques.

Ce guide est principalement conçu pour couvrir les services des catégories de l'IA de classification et de l'apprentissage automatique par IA prédictive.

En outre, AWS propose du matériel spécialisé et accéléré pour un entraînement et une inférence ML de haute performance.

- Les instances [Amazon EC2 P5](#) sont équipées du processeur NVIDIA H100 Tensor Core GPUs, qui convient parfaitement aux tâches de formation et d'inférence dans le cadre de l'apprentissage automatique. Les instances [Amazon EC2 G5](#) comportent jusqu'à 8 processeurs NVIDIA A10G Tensor Core et des processeurs AMD EPYC de deuxième génération GPUs, pour un large éventail de cas d'utilisation gourmands en ressources graphiques et en apprentissage automatique.
- [AWS Trainium](#) est l'accélérateur de machine learning de deuxième génération spécialement AWS conçu pour l'apprentissage en profondeur (DL) de plus de 100 milliards de modèles de paramètres.
- AWS Inferentia Les [instances Amazon EC2 Inf2 basées sur 2](#) sont conçues pour fournir des performances élevées au moindre coût sur Amazon EC2 pour vos applications d'inférence DL et d'IA générative.

Tenez compte

Lors de la résolution d'un problème commercial avec les services de machine AWS learning, la prise en compte de plusieurs critères clés peut contribuer à garantir le succès. La section suivante décrit certains des principaux critères à prendre en compte lors du choix d'un service de machine learning.

Problem definition

Définition du problème

La première étape du cycle de vie du machine learning consiste à définir le problème commercial. Comprendre le problème que vous essayez de résoudre est essentiel pour choisir le bon service de machine AWS learning, car différents services sont conçus pour résoudre différents problèmes. Il est également important de déterminer si le ML convient le mieux au problème de votre entreprise.

Une fois que vous avez déterminé que le ML est le mieux adapté, vous pouvez commencer par choisir parmi une gamme de services d' AWS IA spécialement conçus (dans des domaines tels que la parole, la vision et les documents).

Amazon SageMaker AI fournit une infrastructure entièrement gérée si vous devez créer et entraîner vos propres modèles. AWS propose une gamme de frameworks de ML avancés et de choix d'infrastructures pour les cas où vous avez besoin de modèles de ML hautement personnalisés et spécialisés. AWS propose également un large éventail de modèles de base populaires pour créer de nouvelles applications avec l'IA générative.

ML algorithm

Algorithme ML

Le choix de l'algorithme ML pour le problème métier que vous essayez de résoudre dépend du type de données avec lequel vous travaillez, ainsi que des résultats souhaités. Les informations suivantes décrivent comment chacune des principales catégories de services d' AWS IA/ML vous permet de travailler avec ses algorithmes :

- Services d'IA spécialisés : Ces services offrent une capacité limitée de personnalisation de l'algorithme ML, car il s'agit de modèles pré-entraînés optimisés pour des tâches spécifiques. Vous pouvez généralement personnaliser les données d'entrée et certains paramètres, mais vous n'avez pas accès aux modèles ML sous-jacents ni la possibilité de créer vos propres modèles.
- Amazon SageMaker AI : ce service offre le plus de flexibilité et de contrôle sur l'algorithme ML. Vous pouvez utiliser l' SageMaker IA pour créer des modèles personnalisés à l'aide de vos propres algorithmes et frameworks, ou utiliser des modèles et algorithmes prédéfinis fournis par AWS. Cela permet un degré élevé de personnalisation et de contrôle du processus d'apprentissage automatique.
- Frameworks et infrastructures de ML de niveau inférieur : ces services offrent le plus de flexibilité et de contrôle sur l'algorithme de ML. Vous pouvez utiliser ces services pour créer des modèles de machine learning hautement personnalisés à l'aide de leurs propres algorithmes et frameworks. Cependant, l'utilisation de ces services nécessite une expertise significative en matière de machine learning et peut ne pas être réalisable dans tous les cas d'utilisation.

Security

Sécurité

Si vous avez besoin d'un point de terminaison privé dans votre VPC, vos options varient en fonction de la couche de services AWS ML que vous utilisez. Il s'agit des licences suivantes :

- Services d'IA spécialisés : La plupart des services d'IA spécialisés ne prennent actuellement pas en charge les points de terminaison privés dans VPCs. Toutefois, les étiquettes personnalisées Amazon Rekognition et Amazon Comprehend Custom sont accessibles via des points de terminaison VPC.
- Les principaux services d'intelligence artificielle : Amazon Translate, Amazon Transcribe et Amazon Comprehend prennent tous en charge les points de terminaison VPC.
- Amazon SageMaker AI : SageMaker AI fournit un support intégré pour les points de terminaison VPC, vous permettant de déployer leurs modèles entraînés en tant que point de terminaison accessible uniquement depuis leur VPC.
- Frameworks et infrastructure de machine learning de niveau inférieur : vous pouvez déployer vos modèles sur des EC2 instances Amazon ou dans des conteneurs au sein de votre VPC, afin de contrôler totalement la configuration réseau.

Latency

Latence

Les services d'intelligence artificielle de haut niveau, tels qu'Amazon Rekognition et Amazon Transcribe, sont conçus pour gérer une grande variété de cas d'utilisation et offrir des performances élevées en termes de rapidité. Cependant, il est possible qu'ils ne répondent pas à certaines exigences de latence.

Si vous utilisez des frameworks et une infrastructure de machine learning de niveau inférieur, nous vous recommandons de tirer parti d'Amazon SageMaker AI. Cette option est généralement plus rapide que la création de modèles personnalisés en raison de son service entièrement géré et de ses options de déploiement optimisées. Bien qu'un modèle personnalisé hautement optimisé puisse surpasser l' SageMaker IA, son développement nécessitera une expertise et des ressources importantes.

Accuracy

Précision

La précision des services AWS ML varie en fonction du cas d'utilisation spécifique et du niveau de personnalisation requis. Les services d'intelligence artificielle de haut niveau, tels qu'Amazon

Rekognition, sont basés sur des modèles pré-entraînés qui ont été optimisés pour des tâches spécifiques et offrent une grande précision dans de nombreux cas d'utilisation.

Dans certains cas, vous pouvez choisir d'utiliser Amazon SageMaker AI, qui fournit une plateforme plus flexible et personnalisable pour créer et former des modèles de machine learning personnalisés. En créant vos propres modèles, vous pourrez peut-être obtenir une précision encore plus élevée que ce qui est possible avec des modèles préentraînés.

Vous pouvez également choisir d'utiliser des frameworks et des infrastructures de machine learning, tels qu' TensorFlow Apache MXNet, pour créer des modèles hautement personnalisés offrant la plus grande précision possible pour votre cas d'utilisation spécifique.

AWS and responsible AI

AWS et IA responsable

AWS construit des modèles de base (FMs) en gardant à l'esprit l'IA responsable à chaque étape de son processus de développement. Tout au long de la conception, du développement, du déploiement et des opérations, nous prenons en compte une série de facteurs, notamment :

1. Exactitude (dans quelle mesure un résumé correspond au document sous-jacent ; exactitude factuelle d'une biographie)
2. Équité (si les résultats traitent les groupes démographiques de la même manière)
3. Considérations relatives à la propriété intellectuelle et au copyright
4. Utilisation appropriée (filtrage des demandes de conseils juridiques, de diagnostics médicaux ou d'activités illégales des utilisateurs)
5. Toxicité (discours haineux, blasphèmes et insultes)
6. Confidentialité (protection des informations personnelles et des instructions des clients)

AWS intègre des solutions pour résoudre ces problèmes dans les processus utilisés pour acquérir les données de formation, dans les processus FMs eux-mêmes et dans la technologie utilisée pour prétraiter les instructions des utilisateurs et les résultats après le traitement.

Choix

Maintenant que vous connaissez les critères selon lesquels vous allez évaluer les options de vos services de ML, vous êtes prêt à choisir le service de AWS ML qui convient le mieux aux besoins de

vos organisation. Le tableau suivant indique quels services de machine learning sont optimisés pour quelles circonstances. Utilisez-le pour déterminer le service AWS ML le mieux adapté à votre cas d'utilisation.

Catégories	Quand l'utiliserez-vous ?	Pour quoi est-il optimisé ?	AI/ML Services ou environnements connexes
<p>Cas d'utilisation spécifiques</p> <p>Ces services d'intelligence artificielle sont destinés à répondre à des besoins spécifiques. Ils incluent la personnalisation, les prévisions, la détection des anomalies, la transcription vocale, etc. Comme ils sont fournis sous forme de services, ils peuvent être intégrés dans des applications sans aucune expertise en machine learning.</p>	<p>Utilisez les services d'IA fournis AWS lorsque vous avez besoin de fonctionnalités spécifiques et prédéfinies à intégrer à vos applications, sans avoir besoin de personnalisation approfondies ou d'expertise en apprentissage automatique. Ces services sont conçus pour être faciles à utiliser et ne nécessitent pas beaucoup de codage ou de configuration.</p>	<p>Ces services sont conçus pour être faciles à utiliser et ne nécessitent pas beaucoup d'expertise en matière de codage, de configuration ou d'apprentissage automatique.</p>	<p>Amazon Augmented AI</p> <p>Amazon CodeGuru</p> <p>Amazon Comprehend</p> <p>Amazon Comprehend Medical</p> <p>Amazon DevOps Guru</p> <p>Amazon Forecast</p> <p>Amazon Kendra</p> <p>Amazon Lex</p> <p>Amazon Personalize</p> <p>Amazon Polly</p> <p>Amazon Rekognition</p> <p>Amazon Textract</p> <p>Amazon Transcribe</p> <p>Amazon Translate</p>
<p>Services de machine learning</p>	<p>Utilisez ces services lorsque vous avez</p>	<p>Ces services sont optimisés</p>	<p>Amazon SageMaker AI</p>

Catégories	Quand l'utiliserez-vous ?	Pour quoi est-il optimisé ?	AI/ML Services ou environnements connexes
<p>Ces services peuvent être utilisés pour développer des modèles ou des flux de travail d'apprentissage automatique personnalisés qui vont au-delà des fonctionnalités prédéfinies proposées par les principaux services d'IA.</p>	<p>besoin de modèles d'apprentissage automatique ou de flux de travail plus personnalisés qui vont au-delà des fonctionnalités prédéfinies proposées par les principaux services d'IA.</p>	<p>pour la création et la formation de modèles d'apprentissage automatique personnalisés, la formation à grande échelle sur plusieurs instances ou clusters de GPU, un meilleur contrôle sur le déploiement de modèles d'apprentissage automatique, l'inférence en temps réel et pour la création de end-to-end flux de travail.</p>	<p>Amazon SageMaker AI JumpStart</p> <p>SageMaker Studio d'IA</p> <p>SageMaker Canevas AI</p> <p>SageMaker Laboratoire AI Studio</p> <p>SageMaker IA : Ground Truth</p> <p>PyTorch sur AWS</p> <p>Apache MxNet</p> <p>Hugging Face</p> <p>TensorFlow sur AWS</p>
<p>Infrastructures</p> <p>Pour déployer le machine learning en production, vous avez besoin d'une infrastructure rentable, qu'Amazon met en œuvre avec du silicium AWS intégré.</p>	<p>À utiliser lorsque vous souhaitez réduire le coût des modèles de formation et que vous devez exécuter des inférences dans le cloud.</p>	<p>Optimisé pour soutenir le déploiement rentable de l'apprentissage automatique.</p>	<p>AWS Trainium</p> <p>AWS Inferentia et Inferentia2</p> <p>Amazon SageMaker AI HyperPod</p>

Catégories	Quand l'utiliserez-vous ?	Pour quoi est-il optimisé ?	AI/ML Services ou environnements connexes
<p>Outils et services associés</p> <p>Ces outils et les services associés sont conçus pour faciliter le déploiement de l'apprentissage automatique.</p>	<p>Ces services et outils sont conçus pour vous aider à accélérer le deep learning dans le cloud, en fournissant des images de machine Amazon, des images docker et la résolution d'entités.</p>	<p>Optimisé pour vous aider à accélérer le deep learning dans le cloud.</p>	<p>AWS Apprentissage profond (deep learning) AMIs</p> <p>AWS Conteneurs Deep Learning</p> <p>AWS Résolution de l'entité</p>

Utiliser

Maintenant que vous comprenez clairement les critères à appliquer pour choisir un service de AWS ML, vous pouvez sélectionner le ou les services d' AWS IA/ML optimisés pour les besoins de votre entreprise.

Pour découvrir comment utiliser et en savoir plus sur le ou les services que vous avez choisis, nous avons proposé trois ensembles de parcours pour explorer le fonctionnement de chaque service. Le premier ensemble de parcours fournit une documentation approfondie, des didacticiels pratiques et des ressources pour démarrer avec Amazon Comprehend, Amazon Textract, Amazon Translate, Amazon Lex, Amazon Polly, Amazon Rekognition et Amazon Transcribe.

Amazon Comprehend

- Commencez avec Amazon Comprehend

Utilisez la console Amazon Comprehend pour créer et exécuter une tâche de détection d'entités asynchrone.

[Commencez avec le didacticiel »](#)

- Analysez des informations sous forme de texte avec Amazon Comprehend

Découvrez comment utiliser Amazon Comprehend pour analyser du texte et en tirer des enseignements.

[Commencez avec le didacticiel »](#)

- Tarification d'Amazon Comprehend

Découvrez des informations sur les tarifs d'Amazon Comprehend et des exemples.

[Découvrez le guide »](#)

Amazon Textract

- Commencer à utiliser Amazon Textract

Découvrez comment Amazon Textract peut être utilisé avec du texte mis en forme pour détecter des mots et des lignes de mots proches les uns des autres, ainsi que pour analyser un document à la recherche d'éléments tels que du texte associé, des tableaux, des paires clé-valeur et des éléments de sélection.

[Découvrez le guide »](#)

- Extrayez du texte et des données structurées avec Amazon Textract

Découvrez comment utiliser Amazon Textract pour extraire du texte et des données structurées d'un document.

[Commencez avec le didacticiel »](#)

- AWS Power Hour : Machine Learning

Découvrez Amazon Textract dans cet épisode, passez du temps à le AWS Management Console lire et passez en revue des exemples de code qui vous aideront à comprendre comment tirer le meilleur parti du service. APIs

[Regardez la vidéo »](#)

Amazon Translate

- Commencer à utiliser Amazon Translate à l'aide de la console

Le moyen le plus simple de démarrer avec Amazon Translate est d'utiliser la console pour traduire du texte. Apprenez à traduire jusqu'à 10 000 caractères à l'aide de la console.

[Découvrez le guide »](#)

- Translate du texte d'une langue à l'autre dans le cloud

Dans cet exemple de didacticiel, en tant que membre d'une entreprise internationale de fabrication de bagages, vous devez comprendre ce que les clients disent à propos de votre produit dans les avis rédigés dans la langue du marché local, à savoir le français.

[Commencez avec le didacticiel »](#)

- Tarifs d'Amazon Translate

Découvrez les tarifs d'Amazon Translate, y compris le niveau gratuit, qui fournit 2 millions de caractères par mois pendant 12 mois.

[Découvrez le guide »](#)

Amazon Lex

- Guide du développeur Amazon Lex V2

Découvrez les informations relatives à la mise en route, à son fonctionnement et à la tarification d'Amazon Lex V2.

[Découvrez le guide »](#)

- Présentation d'Amazon Lex Nous vous présentons le service conversationnel Amazon Lex et vous présentons des exemples qui vous montrent comment créer un bot et le déployer sur différents services de chat.

[Suivez le cours »](#) (connexion requise)

- ~~Exploration de l'IA générative dans les expériences conversationnelles~~

Explorez l'utilisation de l'IA générative dans les expériences de conversation.

[Lisez le blog »](#)

Amazon Polly

- Qu'est-ce qu'Amazon Polly ?

Découvrez un aperçu complet du service cloud qui convertit le texte en discours réaliste et peut être utilisé pour développer des applications visant à accroître l'engagement et l'accessibilité de vos clients.

[Découvrez le guide »](#)

- Surlignez le texte tel qu'il est prononcé à l'aide d'Amazon Polly

Nous vous présentons des méthodes permettant de surligner le texte pendant qu'il est prononcé afin d'ajouter des fonctionnalités visuelles au son dans des livres, des sites Web, des blogs et d'autres expériences numériques.

[Lisez le blog »](#)

- Créez de l'audio pour du contenu en plusieurs langues avec le même personnage vocal TTS dans Amazon Polly

Nous expliquons Neural Text-to-Speech (NTTS) et expliquons comment un large portefeuille de voix disponibles, fournissant une gamme de locuteurs distincts dans les langues prises en charge, peut fonctionner pour vous.

[Lisez le blog »](#)

Amazon Rekognition

- Qu'est-ce qu'Amazon Rekognition ?

Découvrez comment utiliser ce service pour ajouter une analyse d'images et de vidéos à vos applications.

[Découvrez le guide »](#)

- Rekognition pratique : analyse automatique des images et des vidéos

Découvrez comment fonctionne la reconnaissance faciale avec le streaming vidéo, ainsi que des exemples de code et des points clés à un rythme autoguidé.

[Commencez avec le didacticiel »](#)

- Amazon Rekognition FAQs

Découvrez les bases d'Amazon Rekognition et découvrez comment il peut vous aider à améliorer votre apprentissage en profondeur et à analyser visuellement vos applications.

[Lisez le FAQs »](#)

Amazon Transcribe

- Qu'est-ce qu'Amazon Transcribe ?

Découvrez le service de reconnaissance vocale AWS automatique utilisant le ML pour convertir le son en texte. Découvrez comment utiliser ce service en tant que transcription autonome ou comment ajouter des speech-to-text fonctionnalités à n'importe quelle application.

[Découvrez le guide »](#)

- Tarification d'Amazon Transcribe

Nous vous présentons la AWS pay-as-you-go transcription, y compris les options de modèles linguistiques personnalisés et le niveau gratuit d'Amazon Transcribe.

[Découvrez le guide »](#)

- Créez une transcription audio avec Amazon Transcribe

Découvrez comment utiliser Amazon Transcribe pour créer une transcription texte de fichiers audio enregistrés à l'aide d'un scénario d'utilisation réel afin de les tester en fonction de vos besoins.

[Commencez avec le didacticiel »](#)

- Créez une application de streaming Amazon Transcribe

Découvrez comment créer une application pour enregistrer, transcrire et traduire du son en direct en temps réel, les résultats vous étant envoyés directement par e-mail.

[Découvrez le guide »](#)

Le deuxième ensemble de Service AWS parcours IA/ML fournit une documentation approfondie, des didacticiels pratiques et des ressources pour démarrer avec les services de la famille Amazon SageMaker AI.

SageMaker AI

- Comment fonctionne Amazon SageMaker AI

Découvrez un aperçu de l'apprentissage automatique et du fonctionnement de SageMaker l'IA.

[Découvrez le guide »](#)

- Commencer à utiliser Amazon SageMaker AI

Découvrez comment rejoindre un domaine Amazon SageMaker AI pour accéder à Amazon SageMaker AI Studio et à RStudio l' SageMaker IA.

[Découvrez le guide »](#)

- Utiliser Apache Spark avec Amazon SageMaker AI

Découvrez comment utiliser Apache Spark pour le prétraitement des données et l' SageMaker IA pour la formation et l'hébergement des modèles.

[Découvrez le guide »](#)

- Utiliser des conteneurs Docker pour créer des modèles

Découvrez comment Amazon SageMaker AI utilise largement les conteneurs Docker pour les tâches de création et d'exécution. Découvrez comment déployer les images Docker prédéfinies

pour ses algorithmes intégrés et les frameworks d'apprentissage profond pris en charge utilisés pour la formation et l'inférence.

[Découvrez le guide »](#)

- Cadres et langages d'apprentissage automatique

Découvrez comment démarrer avec l' SageMaker IA à l'aide du SDK Amazon SageMaker AI Python.

[Découvrez le guide »](#)

SageMaker AI Autopilot

- Créez une expérience Amazon SageMaker AI Autopilot pour les données tabulaires

Découvrez comment créer une expérience Amazon SageMaker AI Autopilot pour explorer, prétraiter et former différents modèles candidats sur un jeu de données tabulaire.

[Découvrez le guide »](#)

- Création automatique de modèles d'apprentissage automatique

Découvrez comment utiliser Amazon SageMaker AI Autopilot pour créer, entraîner et régler automatiquement un modèle de machine learning, et comment déployer le modèle pour établir des prédictions.

[Commencez avec le didacticiel »](#)

- Découvrez la modélisation avec Amazon SageMaker AI Autopilot avec ces exemples de blocs-notes

Découvrez des exemples de blocs-notes pour le marketing direct, la prévision du taux de désabonnement des clients et découvrez comment intégrer votre propre code de traitement des données à Amazon SageMaker AI Autopilot.

[Découvrez le guide »](#)

SageMaker AI Canvas

- Commencez à utiliser Amazon SageMaker AI Canvas

Découvrez comment commencer à utiliser SageMaker AI Canvas.

[Découvrez le guide »](#)

- Générez des prédictions de machine learning sans écrire de code

Ce didacticiel explique comment utiliser Amazon SageMaker AI Canvas pour créer des modèles de machine learning et générer des prédictions précises sans écrire une seule ligne de code.

[Commencez avec le didacticiel »](#)

- Approfondissez les connaissances d' SageMaker AI Canvas

Découvrez un aperçu détaillé d' SageMaker AI Canvas et de ses fonctionnalités d'apprentissage automatique visuelles sans code.

[Lisez le blog »](#)

- Utilisez Amazon SageMaker AI Canvas pour créer votre premier modèle ML

Découvrez comment utiliser Amazon SageMaker AI Canvas pour créer un modèle de machine learning afin d'évaluer la fidélisation des clients, sur la base d'une campagne par e-mail pour de nouveaux produits et services.

[Commencez par le laboratoire »](#)

SageMaker AI Data Wrangler

- Commencer à utiliser Amazon SageMaker AI Data Wrangler

Découvrez comment configurer SageMaker AI Data Wrangler, puis fournissez une procédure pas à pas à l'aide d'un exemple de jeu de données existant.

[Découvrez le guide »](#)

Découvrez comment préparer des données pour le machine learning à l'aide d'Amazon SageMaker AI Data Wrangler.

[Commencez avec le didacticiel »](#)

- SageMaker Atelier de plongée approfondie sur AI Data Wrangler

Apprenez à appliquer les types d'analyse appropriés à votre jeu de données pour détecter les anomalies et les problèmes, results/insights à utiliser le dérivé pour formuler des actions correctives lors des transformations de votre ensemble de données et à tester le bon choix et la bonne séquence de transformations à l'aide des options de modélisation rapide fournies par SageMaker AI Data Wrangler.

[Commencez par l'atelier »](#)

SageMaker AI Ground Truth

- Commencer à utiliser Amazon Ground Truth

Découvrez comment utiliser la console pour créer une tâche d'étiquetage, affecter une main-d'œuvre publique ou privée et envoyer la tâche d'étiquetage à votre personnel. Découvrez comment suivre la progression d'une tâche d'étiquetage.

[Découvrez le guide »](#)

- Étiqueter les données d'entraînement pour le Machine Learning

Découvrez comment configurer une tâche d'étiquetage dans Amazon SageMaker AI Ground Truth afin d'annoter les données d'entraînement pour votre modèle de machine learning.

[Commencez avec le didacticiel »](#)

- Débuter avec Amazon Ground Truth Plus Découvrez comment effectuer les étapes nécessaires pour démarrer un projet Amazon SageMaker AI Ground Truth Plus, vérifier les labels et satisfaire aux exigences d' Amazon SageMaker AI Ground Truth Plus.

[Découvrez le guide »](#)

- Commencez avec Amazon Ground Truth Regardez comment commencer à étiqueter vos données en quelques minutes grâce à la console SageMaker AI Ground Truth.

[Regardez la vidéo »](#)

- Amazon SageMaker AI Ground Truth Plus : créez des ensembles de données de formation sans code ni ressources internes

Découvrez Ground Truth Plus, un service clé en main qui fait appel à une main-d'œuvre experte pour fournir rapidement des ensembles de données de formation de haute qualité et qui réduit les coûts jusqu'à 40 %.

[Lisez le blog »](#)

SageMaker AI JumpStart

- Initiez-vous à l'apprentissage automatique grâce à l' SageMaker IA JumpStart

Découvrez des modèles de JumpStart solutions d' SageMaker IA qui configurent une infrastructure pour les cas d'utilisation courants, ainsi que des exemples de blocs-notes exécutables pour l'apprentissage automatique avec l' SageMaker IA.

[Découvrez le guide »](#)

- Commencez rapidement votre projet d'apprentissage automatique à l'aide d'Amazon SageMaker AI JumpStart

Découvrez comment accélérer votre projet de machine learning à l'aide de modèles préformés et de solutions prédéfinies proposées par Amazon AI. SageMaker JumpStart Vous pouvez ensuite déployer le modèle sélectionné via les blocs-notes Amazon SageMaker AI Studio.

[Commencez avec le didacticiel »](#)

- Découvrez Amazon SageMaker AI JumpStart grâce à cet atelier d'une journée d'immersion

Découvrez comment les fonctionnalités d'apprentissage automatique à faible code présentes dans Amazon SageMaker AI Data Wrangler, Autopilot et Jumpstart permettent d'expérimenter plus rapidement et de mettre en production des modèles très précis.

[Commencez par l'atelier »](#)

SageMaker AI Pipelines

- Commencer à utiliser Amazon SageMaker AI Pipelines

Découvrez comment créer des end-to-end flux de travail qui gèrent et déploient des tâches liées à l' SageMaker IA. SageMaker AI Pipelines intègre le SDK SageMaker AI Python, ce qui vous permet de créer chaque étape de votre pipeline à l'aide d'une interface basée sur Python.

[Découvrez le guide »](#)

- Automatisez les flux de travail d'apprentissage

Découvrez comment créer et automatiser des flux de travail d'apprentissage end-to-end automatique (ML) à l'aide d'Amazon SageMaker AI Pipelines, d'Amazon SageMaker AI Model Registry et d'Amazon SageMaker AI Clarify.

[Commencez avec le didacticiel »](#)

- Comment créer des flux de travail ML entièrement automatisés avec Amazon SageMaker AI Pipelines

Découvrez Amazon SageMaker AI Pipelines, le premier CI/CD service de machine learning au monde conçu pour être accessible à tous les développeurs et data scientists. SageMaker AI Pipelines intègre des CI/CD pipelines au ML, réduisant ainsi le temps de codage requis.

[Regardez la vidéo »](#)

SageMaker AI Studio

- Créez et entraînez un modèle d'apprentissage automatique localement

Découvrez comment créer et entraîner un modèle de machine learning localement dans votre bloc-notes Amazon SageMaker AI Studio.

[Commencez avec le didacticiel »](#)

- SageMaker Intégration d'AI Studio à l'atelier EMR

Découvrez comment utiliser le traitement distribué à grande échelle pour préparer les données, puis entraîner des modèles d'apprentissage automatique.

[Commencez par l'atelier »](#)

Le troisième ensemble de Service AWS parcours d'IA/ML fournit une documentation approfondie, des didacticiels pratiques et des ressources pour démarrer avec AWS Trainium Amazon AWS Inferentia Titan.

AWS Trainium

- Élargissement de la formation distribuée avec AWS Trainium Amazon EKS

Découvrez comment tirer parti de la disponibilité générale des instances Amazon EC2 Trn1 grâce à AWS Trainium un accélérateur de machine learning spécialement conçu pour fournir une plateforme performante, rentable et extrêmement évolutive pour la formation de modèles d'apprentissage profond dans le cloud.

[Lisez le blog »](#)

- Vue d'ensemble de AWS Trainium

Découvrez AWS Trainium l'accélérateur d'apprentissage automatique (ML) de deuxième génération AWS spécialement conçu pour l'apprentissage en profondeur de plus de 100 milliards de modèles de paramètres. Chaque instance Amazon Elastic Compute Cloud (EC2) Trn1 déploie jusqu'à 16 AWS Trainium accélérateurs afin de fournir une solution performante et peu coûteuse pour la formation au deep learning (DL) dans le cloud.

[Découvrez le guide »](#)

- Instances Trainium recommandées

Découvrez comment les AWS Trainium instances sont conçues pour fournir des performances élevées et une rentabilité élevées pour les charges de travail d'inférence de modèles de deep learning.

[Découvrez le guide »](#)

AWS Inferentia

- Vue d'ensemble de AWS Inferentia

Découvrez comment les accélérateurs sont conçus AWS pour fournir des performances élevées au moindre coût à vos applications d'inférence basées sur le deep learning (DL).

[Découvrez le guide »](#)

- AWS Inferentia 2 s'appuie sur AWS Inferentia 1 en offrant un débit 4 fois plus élevé et une latence 10 fois plus faible

Découvrez pour quoi AWS Inferentia 2 est optimisé et découvrez comment il a été conçu dès le départ pour offrir de meilleures performances tout en réduisant le coût LLMs et en réduisant les inférences génératives de l'IA.

[Lisez le blog »](#)

- Inférence par apprentissage automatique à l'aide de AWS Inferentia

Découvrez comment créer un cluster Amazon EKS avec des nœuds exécutant des instances Amazon EC2 Inf1 et (éventuellement) déployer un exemple d'application. Les instances Amazon EC2 Inf1 sont alimentées par des AWS Inferentia puces conçues sur mesure AWS pour fournir des performances élevées et une inférence à moindre coût dans le cloud.

[Découvrez le guide »](#)

Amazon Titan

- Présentation d'Amazon Titan

Découvrez comment Amazon Titan FMs est préentraîné sur de grands ensembles de données, ce qui en fait de puissants modèles polyvalents. Découvrez comment vous pouvez les utiliser tels quels, ou en privé, pour les personnaliser avec vos propres données pour une tâche particulière sans annoter de gros volumes de données.

[Découvrez le guide »](#)

Explorez

- Schémas d'architecture

Ces diagrammes d'architecture de référence présentent des exemples de services d' AWS intelligence artificielle et de machine learning utilisés.

[Explorez les diagrammes d'architecture »](#)

- Livres blancs

Consultez les livres blancs pour vous aider à démarrer et à découvrir les meilleures pratiques en matière de choix et d'utilisation AI/ML des services.

[Découvrez les livres blancs »](#)

- AWS Des solutions

Découvrez des solutions approuvées et des conseils architecturaux pour les cas d'utilisation courants des services d'intelligence artificielle et de machine learning.

[Découvrez les solutions »](#)

Ressources

Modèles de fondation

Les modèles de fondation pris en charge incluent :

- [Claude anthropique](#)
- [Cohérence de la commande et intégration](#)
- [AI21 Laboratoires Jurassic](#)
- [Méta lama](#)
- [IA Mistral](#)
- [Diffusion stable XL](#)
- [Amazon Titan](#)

À l'aide d'Amazon Bedrock, vous pouvez tester différents modèles de base et les personnaliser en privé avec vos données.

Cas d'utilisation ou services spécifiques au secteur

- [Amazon Comprehend Medical](#)
- [Amazon Fraud Detector](#)
- [AWS HealthLake](#)
- [Amazon Lookout for Equipment](#)
- [Amazon Lookout for Metrics](#)
- [Amazon Lookout for Vision](#)
- [Amazon Monitron](#)
- [AWS HealthOmics](#)
- [AWS Panorama](#)

Articles de blog associés

- [De nouvelles fonctionnalités importantes facilitent l'utilisation d'Amazon Bedrock pour créer et faire évoluer des applications d'IA génératives, et permettent d'obtenir des résultats impressionnants](#)
- [AWS Déterminez et AWS Trainium réduisez les coûts de déploiement des modèles Llama 3 dans Amazon AI SageMaker JumpStart](#)
- [Révolutionnez la satisfaction client grâce à des modèles de récompenses personnalisés pour votre entreprise sur Amazon AI SageMaker](#)
- [Amazon Personalize lance de nouvelles recettes compatibles avec des catalogues d'articles plus volumineux avec une latence plus faible](#)

Historique du document

Le tableau suivant décrit les modifications importantes apportées à ce guide de décision. Pour recevoir des notifications concernant les mises à jour de ce guide, vous pouvez vous abonner à un flux RSS.

Modification	Description	Date
Mise à jour mineure	Contenu mis à jour pour Amazon Q et les derniers outils d'intelligence artificielle et de machine learning d'Amazon.	3 mai 2024
Première version	Publication initiale du guide de décision.	24 juillet 2023

Les traductions sont fournies par des outils de traduction automatique. En cas de conflit entre le contenu d'une traduction et celui de la version originale en anglais, la version anglaise prévaudra.